Perturbative quantum gauge invariance: Where the ghosts come from.

Andreas Aste
Institute for Theoretical Physics
Klingelbergstrasse 82, Basel, Switzerland

January 23, 2003

Abstract

A condensed introduction to quantum gauge theories is given in the perturbative S-matrix framework; path integral methods are used nowhere. This approach emphasizes the fact that it is not necessary to start from classical gauge theories which are then subject to quantization, but it is also possible to recover the classical group structure and coupling properties from purely quantum mechanical principles. As a main tool we use a free field version of the Becchi-Rouet-Stora-Tyutin gauge transformation, which contains no interaction terms related to a coupling constant. This free gauge transformation can be formulated in an analogous way for quantum electrodynamics, Yang-Mills theories with massless or massive gauge bosons and quantum gravity.

Keywords: Quantum gauge field theory, ghosts.

PACS: 11.10.-z, 11.15.Bt, 12.20.Ds, 12.38.Bx.

1 Introduction

It is well known that gauge theories play a fundamental role in modern quantum field theory. From this observation one might conclude that gauge invariance is a physical mechanism inherent in many particle interactions. But this is not the case. Gauge invariance is rather a mathematical artefact which stems from the way we formulate (quantum) field theory, although its presence is very helpful in order to achieve a consistent formulation of many quantum field theories. Like in the case of general relativity, where we are free to choose the coordinate system according to our taste, we also have some freedom how to deal with field theories. The free quantum fields which are used in perturbation theory in order to calculate physical quantities like e.g. cross sections simply provide a 'coordinatization' of the problem under consideration, and they are 'as physical' as a coordinate system.

One may illustrate this by a simple observation. Massless particles with spin have only two degrees of freedom, e.g. photons appear only in two helicity states. Such states are called 'physical' in the paper, whereas timelike or longitudinal photons, which are an unavoidable byproduct of covariant quantization, are called 'unphysical'. But if we take into account that one can introduce running coupling constants and running masses in interacting field theories, it becomes clear that also 'physical' particles like quarks are merely mathematical constructions, since we are always free to change the structure of the Fock space underlying our calculations by renormalizing the mass of the quark asymptotic states. Therefore it is meaningless to ask how a composite particle can be decomposed into free 'naked' particle states with a well-defined mass. What one really should do is to calculate vacuum expectation values of products of local interacting fields in a nonperturbative manner. From these distributions it is in principle possible to reconstruct the true physical Hilbert space, as it is claimed by the reconstruction theorem of Wightman [1]. But this is a very difficult task and has been performed only for exotic cases like e.g. quantum field theories in 1+1 spacetime dimensions [2, 3, 4, 5].

It is not the intention of this introduction to give a complete description of all aspects related to the arbitraryness of the formulation of quantum field theory, and for mathematical details we must refer to the literature. But since we are working in a strictly perturbative sense by using free field operators only, the whole formalism used in this paper can be formulated in a fully mathematical manner if necessary.

The structure of this paper is as follows: First, we start with the free classical electromagnetic field (or, more precisely, with the classical gauge theory of a massless spin 1 field), for which the basic issues of classical gauge invariance are explained. Then the electromagnetic field is subject to quantization, and we lift the gauge transformation on the quantum level. It is explained how the quantum gauge transformation, which is the free field analogue of the full Becchi-Rouet-Stora-Tyutin (BRST) transformation [6, 7], can be applied to quantum chromodynamics (QCD) along the same lines as for the quantum electrodynamics (QED) case. As a further step, the method is extended to massive gauge theories, where the Higgs field is involved. Finally, we outline how quantum gauge invariance can be formulated for quantum gravity.

2 The classical electromagnetic field

The equation of motion for the noninteracting vector potential A^μ in electrodynamics can be derived from the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = \frac{1}{2}(\vec{E}^2 - \vec{B}^2), \tag{1}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength tensor which can be expressed directly in terms of the components of the electric and magnetic fields $\vec{E}$ and $\vec{B}$. From the Euler-Lagrange equations

$$\partial_\nu \frac{\delta \mathcal{L}}{\delta \partial_\nu A_\mu} - \frac{\delta \mathcal{L}}{\delta A_\mu} = 0 \tag{2}$$

one derives then the wave equation

$$\partial_\nu \partial^\nu A_\mu - \partial_\mu \partial_\nu A^\nu = 0. \tag{3}$$

A problem arises from the fact that the physically measurable electromagnetic fields $\vec{E}$ and $\vec{B}$ remain unchanged if the vector potential is subject to a gauge transformation expressed by a real scalar field u ($x = (x^0, \vec{x})$)

$$A_\mu(x) \to A_\mu(x) + \partial_\mu u(x), \tag{4}$$

i.e. there is a redundance in the local description of the electromagnetic fields by a vector potential, aside from global topological information which may be contained in the potential like in the case described by Aharonov and Bohm [8]. From the mathematical point of view, we generally assume that the objects we are dealing with 'behave well', i.e. A_μ and u should be differentiable and behave at large spatial distances in such a way that physical quantities like the total energy of the field configuration are finite.

The possibility to transform the vector potential without changing its physical content allows to impose *gauge conditions*. Due to its manifestly covariant form, a very popular choice is the Lorentz gauge condition, $\partial_\mu A^\mu = 0$. This condition can be enforced by transforming the initial vector potential according to (4) with a field u which fulfills

$$\Box u = \partial_\mu \partial^\mu u = -\partial_\mu A^\mu. \tag{5}$$

Then the new transformed field satisfies the simple wave equation

$$\Box A^\mu = \partial_\nu \partial^\nu A^\mu = 0. \tag{6}$$

But the vector potential is not completely determined by the Lorentz gauge condition. It is still possible to apply a gauge transformation to the vector potential if u fulfills the wave equation $\partial_\nu \partial^\nu u = 0$, such that the transformed vector potential is still in Lorentz gauge. In the sequel we will denote the field u, which is somehow related to the unphysical degrees of freedom of the vector potential in Lorentz gauge, as 'ghost field'. The reason will become clear at a later stage.

It is important to note that the equation of motion for the vector potential can be manipulated by a modification of the Lagrangian according to an early observation by Fermi [9]. Adding a *gauge fixing term* to the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\xi(\partial_\mu A^\mu)^2, \tag{7}$$

where ξ is an arbitrary real parameter, leads to the equation of motion

$$\Box A_\mu - (1-\xi)\partial_\mu\partial_\nu A^\nu = 0. \tag{8}$$

One may argue that the gauge fixing does not change the physical content of the Lagrangian, since we may enforce the Lorentz gauge condition $\partial_\mu A^\mu = 0$, and then the gauge fixing term vanishes in the Lagrangian. For the special choice $\xi = 1$ we recover the wave equation (6). This choice is often referred to as *Feynman gauge*, although this terminology is a bit misleading. We must clearly distinguish between the gauge fixing, which fixes the Lagrangian and determines the form of the wave equation (e.g. the Feynman gauge), and the choice of a gauge condition for the vector potential (e.g. the Lorentz gauge), which reduces the redundance originating from the use of a vector potential as a basic field. It is possible to impose more general (nonlinear) gauge fixing conditions, e.g. by adding a term $\sim (\partial_\mu A^\mu)^4$ to the Lagrangian, but we restrict ourselves to the most important cases here. We conclude this discussion of the gauge fixing with the remark that it is indeed necessary to add appropriate gauge fixing terms to the free Lagrangian (1), since otherwise problems arise in the construction of the free photon propagator.

Up to now, the discussion of the electromagnetic field has been completely classical. In the next section, we will consider a quantized version of the free electromagnetic field. For the field operator $A_\mu(x)$ we will use the Feynman gauge such that the operator obeys the simplest equation of motion (6). All calculations could be performed in a strictly analogous way for arbitrary $\xi-$gauges, but the choice $\xi = 1$ has obviously notational advantages [10]. Strictly speaking, each choice of ξ leads after quantization to a different quantum field theory with different Fock spaces and field operators. But physical observables like cross sections etc. should be independent of the gauge fixing.

3 Quantization of free fields

It is well known that a free massless hermitian scalar field $\varphi(x)$ can be represented by

$$\varphi(t,\vec{x}) = (2\pi)^{-3/2}\int \frac{d^3k}{\sqrt{2\omega}}\Big(a(\vec{k})e^{-i(\omega t-\vec{k}\vec{x})} + a(\vec{k})^\dagger e^{i(\omega t-\vec{k}\vec{x})}\Big), \tag{9}$$

where $\omega = |\vec{k}|$ and the creation and annihilation operators fulfill the commutation relations

$$[a(\vec{k}), a(\vec{k}')^\dagger] = \delta^{(3)}(\vec{k}-\vec{k}'), \tag{10}$$

$$[a(\vec{k}), a(\vec{k}')] = [a(\vec{k})^\dagger, a(\vec{k}')^\dagger] = 0. \tag{11}$$

The † denotes the adjoint with respect to a positive definite scalar product so that the operators can be represented in the usual way in a Fock space with unique vacuum, $\delta^{(3)}$ is the three dimensional Dirac distribution.

We now try to quantize the vector potential $A^\mu(x)$ as four independent real scalar fields

$$A^\mu(t,\vec{x}) = (2\pi)^{-3/2} \int \frac{d^3k}{\sqrt{2\omega}} \Big(a^\mu(\vec{k}) e^{-i(\omega t - \vec{k}\vec{x})} + a^\mu(\vec{k})^\dagger e^{i(\omega t - \vec{k}\vec{x})} \Big), \tag{12}$$

and naively we assume that the commutation relations for the creation and annihilation operators are

$$[a^\mu(\vec{k}), a^\nu(\vec{k}')^\dagger] = \delta_{\mu\nu}\delta^{(3)}(\vec{k}-\vec{k}'),$$

$$[a^\mu(\vec{k}), a^\nu(\vec{k}')] = [a^\mu(\vec{k})^\dagger, a^\nu(\vec{k}')^\dagger] = 0. \tag{13}$$

$\Box A^\mu = 0$ is automatically satisfied according to (12), since the spacetime dependence of the field operator is given by the plane wave terms $e^{\pm ikx} = e^{\pm i(\omega t - \vec{k}\vec{x})}$, and δ is the Kronecker delta of the indices μ, ν.

A calculation shows that the free massless hermitian scalar field φ fulfills the commutation relation

$$[\varphi(x), \varphi(y)] = -iD(x-y), \tag{14}$$

where $D(x-y)$ is the (massless) Pauli-Jordan distribution

$$D(x) = \frac{i}{(2\pi)^3} \int d^4k\, \delta(k^2)\mathrm{sgn}(k^0) e^{-ikx} = \frac{1}{2\pi}\mathrm{sgn}(x^0)\delta(x^2). \tag{15}$$

The Pauli-Jordan distribution fulfills the distributional identity

$$\partial_0 D(x)|_{x_0=0} = \delta^{(3)}(\vec{x}) \tag{16}$$

which implies the equal time commutation relation for a scalar field φ and its canonical momentum $\pi = \dot{\varphi}$

$$[\varphi(x), \dot{\varphi}(y)]|_{x_0=y_0} = i\delta^{(3)}(\vec{x}-\vec{y}) \tag{17}$$

which is often taken as the starting point for the quantization of the scalar field. The distribution $D(x-y)$ was introduced by Pauli and Jordan in 1928, when they performed the first covariant quantization of the radiation field [11]. The naive choice (12) comprises a problem, since the commutation relations

$$[A^\mu(x), A^\nu(y)] = -i\delta_{\mu\nu} D(x-y) \tag{18}$$

are not covariant. A simple way to remedy this defect is to change the definition of A^0 into

$$A^0(x) = (2\pi)^{-3/2} \int \frac{d^3k}{\sqrt{2\omega}} \Big(a^0(\vec{k}) e^{-ikx} - a^0(\vec{k})^\dagger e^{ikx} \Big), \tag{19}$$

i.e. we make A^0 a skew-adjoint operator instead of self-adjoint. Another strategy is to introduce a Fock space with negative norm states, as proposed by Gupta and Bleuler [12, 13]. The commutation relations then become

$$[A^\mu(x), A^\nu(y)] = ig^{\mu\nu} D(x-y). \tag{20}$$

We have introduced a nonhermitian field in order to save the Lorentz symmetry of the theory, and therefore one might expect problems with the unitarity of the QED S-matrix. But no problems arise, a fact which is related to the gauge symmetry of the theory.

A concluding remark is in order here. It is possible to avoid unphysical polarization states by quantizing the photon field in radiation gauge, which is expressed classically by the conditions

$$A^0 = 0, \quad \vec{\nabla} \cdot \vec{A} = 0. \tag{21}$$

The (quantized) photon field operator contains then only physical (transverse) polarizations

$$A^0 = 0, \quad \vec{A} = (2\pi)^{-3/2} \int \frac{d^3k}{\sqrt{2\omega}} \sum_{\lambda=1,2} \vec{\epsilon}(\vec{k},\lambda)[a(\vec{k},\lambda)e^{-ikx} + a^\dagger(\vec{k},\lambda)e^{ikx}] \tag{22}$$

with

$$\vec{\epsilon}(\vec{k},\lambda) \cdot \vec{\epsilon}(\vec{k},\lambda') = \delta_{\lambda\lambda'} \quad \text{and} \quad \vec{\epsilon}(\vec{k},\lambda) \cdot \vec{k} = 0 \tag{23}$$

and standard commutation relations for a and $a^\dagger$. But this is only seemingly an advantage because manifest Lorentz symmetry is lost in (22), and no general strategy is known how to deal with the renormalization of divergent higher order contributions in interacting theories without the guiding help of Lorentz symmetry. Thus we have to make the choice between having unphysical particles, but also manifest Lorentz covariance in our calculations, or working only on a physical Fock space, but without having a consistent scheme at hand to regularize loop diagrams in perturbative calculations.

4 A simple version of a quantum gauge transformation

We introduce now a quantized version of the classical gauge transformation for the free vector potential

$$A^\mu \rightarrow A^\mu + \partial^\mu u, \quad \Box u = 0. \tag{24}$$

Of course there exist different approaches for the treatment of gauge invariance in quantum field theory, but the approach presented here is quite simple and does the job.

As a first step, we define the gauge transformation operator or *gauge charge* Q

$$Q = \int\limits_{x_0=const.} d^3x \, \partial_\mu A^\mu(x) \overleftrightarrow{\partial_0} u(x). \tag{25}$$

It is sufficient for the moment to consider u as a real C-number field. In the case of QCD, u will necessarily become a fermionic scalar field. It can be shown that Q is a well-defined operator on the Fock space generated by the creation and annihilation operators according to (13). It is not important over which spacelike plane the integral in (25) is taken, since Q is time independent:

$$\begin{aligned} \dot{Q} &= \int\limits_{x_0=const.} d^3x \, (-\partial_0^2 \partial_\mu A^\mu u + \partial_\mu A^\mu \partial_0^2 u) \\ &= \int\limits_{x_0=const.} d^3x \, (- \triangle \, \partial_\mu A^\mu u + \partial_\mu A^\mu \triangle \, u) = 0. \end{aligned} \tag{26}$$

This formal proof uses the wave equation and partial integration. Another way to show the time independence of the gauge charge is to define the *gauge current*

$$j_g^\mu = \partial_\nu A^\nu \overleftrightarrow{\partial^\mu} u, \quad Q = \int d^3x \, j_g^0, \tag{27}$$

which is conserved

$$\partial_\mu j_g^\mu = \partial_\mu(\partial_\nu A^\nu \partial^\mu u - \partial^\mu \partial_\nu A^\nu u) = 0. \tag{28}$$

A crucial property of the gauge charge is expressed by the commutator with A^μ, which is a C-number

$$[Q, A^\mu(x)] = i\partial^\mu u(x), \tag{29}$$

and all higher commutators like

$$[Q, [Q, A^\mu(x)]] = 0. \tag{30}$$

vanish for obvious reasons. The reader is invited to check the commutation relation (29), an outline of the calculation can be found in Appendix A. Therefore we have

$$e^{-i\lambda Q} A^\mu e^{+i\lambda Q} = A^\mu - \frac{i\lambda}{1!}[Q, A^\mu] - \frac{\lambda^2}{2!}[Q, [Q, A^\mu]] + ...$$

$$= A^\mu - i\lambda[Q, A^\mu] = A^\mu + \lambda\partial^\mu u, \tag{31}$$

i.e. Q is a *generator* of gauge transformations.

5 Definition of perturbative quantum gauge invariance

We take the next step towards full QED and couple photons to electrons. In perturbative QED, the S-matrix is expanded as a power series in the coupling constant e. At first order, the interaction is described by the normally ordered product of free fields

$$\mathcal{H}_{\text{int}}(x) = -e : \bar{\Psi}(x)\gamma^\mu \Psi(x) : A_\mu(x), \tag{32}$$

where Ψ is the electron field operator. The S-matrix is then usually given in the literature by the formal expression (T denotes time ordering)

$$S = \mathbf{1} + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \ldots d^4x_n \, T[\mathcal{H}_{\text{int}}(x_1) \ldots \mathcal{H}_{\text{int}}(x_n)]$$

$$= \mathbf{1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int d^4x_1 \ldots d^4x_n \, T_n(x_1, \ldots x_n), \tag{33}$$

where we have introduced the time-ordered products T_n for notational simplicity, and we have

$$T_1(x) = -i\mathcal{H}_{\text{int}}(x) = ie : \bar{\Psi}(x)\gamma^\mu \Psi(x) : A_\mu(x). \tag{34}$$

Expression (33) is plagued by infrared and ultraviolet divergences (see Appendix E). We leave this technical problem aside and we assume that the T_n are regularized, well-defined operator valued distributions, which are symmetric in the space coordinates $(x_1, \ldots x_n)$.

In the previous section, we used a real C-number field u in order to show how the gauge transformation can be lifted on the operator level. We assume now that u is a fermionic scalar field, and additionally we introduce an anti-ghost field $\tilde{u}$, which is *not* the hermitian conjugate of u. The reason why we choose u, $\tilde{u}$ to be fermionic, against all conventional wisdom that particles without spin are bosons, is very simple: Bosonic ghosts do not allow to formulate a consistent theory in the case of quantum chromodynamics, and therefore we start directly with the 'correct' strategy. A second reason is that we want Q to be nilpotent; $Q^2 = 0$ which allows the definition of the physical Hilbert space as

$$\mathcal{F}_{phys} = ker\, Q / ran\, Q \tag{35}$$

(see Appendix C).

We give now a precise definition of perturbative quantum gauge invariance for QED, which will work in a completely analogous way for QCD.

Let

$$Q := \int d^3x\, \partial_\mu A^\mu(x) \overleftrightarrow{\partial_0}\, u(x) \tag{36}$$

be the generator of (free field) gauge transformations, called gauge charge for brevity. This Q has first been introduced in a famous paper by Kugo and Ojima [14]. The positive and negative frequency parts of the free fields satisfy the {anti-}commutation relations

$$[A_\mu^{(\pm)}(x), A_\nu^{(\mp)}(y)] = i g_{\mu\nu} D^{(\mp)}(x-y), \tag{37}$$

$$\{u^{(\pm)}(x), \tilde{u}^{(\mp)}(y)\} = -iD^{(\mp)}(x-y), \tag{38}$$

and all other commutators vanish. Here, $D^{\mp}$ are the positive and negative frequency parts of the massless Pauli-Jordan distribution with Fourier transforms

$$\hat{D}^{(\pm)}(p) = \pm \frac{i}{2\pi} \Theta(\pm p^0) \delta(p^2). \tag{39}$$

All the fields fulfill the Klein-Gordon equation with zero mass, and as already mentioned, we are working in Feynman gauge, but the following discussion would go through with some technical changes in other covariant ξ-gauges as well. Note that the quantum field A^μ does not satisfy $\partial_\mu A^\mu = 0$ on the whole Fock space, but only on the physical subspace. An explicit representation of the ghost fields is given by

$$u(x) = (2\pi)^{-3/2} \int \frac{d^3k}{\sqrt{2\omega}} \Big(c_2(\vec{k}) e^{-ikx} + c_1^\dagger(\vec{k}) e^{ikx} \Big), \tag{40}$$

$$\tilde{u}(x) = (2\pi)^{-3/2} \int \frac{d^3k}{\sqrt{2\omega}} \Big(-c_1(\vec{k}) e^{-ikx} + c_2^\dagger(\vec{k}) e^{ikx} \Big), \tag{41}$$

$$\{c_i(\vec{k}), c_j^\dagger(\vec{k}')\} = \delta_{ij} \delta^{(3)}(\vec{k} - \vec{k}'), \quad i,j = 1,2, \tag{42}$$

In order to see how the infinitesimal gauge transformation acts on the free fields, we have to calculate the commutators (see [15] and Appendix A)

$$[Q, A_\mu] = i\partial_\mu u, \quad \{Q, u\} = 0, \quad \{Q, \tilde{u}\} = -i\partial_\nu A^\nu, \quad [Q, \Psi] = [Q, \bar{\Psi}] = 0. \tag{43}$$

The commutators of Q with the electron field are of course trivial, since the operators act on different Fock space sectors. We need only the first and the last commutator in (43) here, the others will become important in the QCD case. From (31) we know that the commutator of Q with an operator gives the first order variation of the operator subject to a gauge transformation. Then we have for the first order interaction T_1

$$[Q, T_1(x)] = -e : \bar{\Psi}(x)\gamma^\mu\Psi(x) : \partial_\mu u(x)$$

$$= i\partial_\mu(ie : \bar{\Psi}(x)\gamma^\mu\Psi(x) : u(x)) = i\partial_\mu T^\mu_{1/1}(x). \tag{44}$$

Here we used the fact that the electron current is conserved

$$\partial_\mu : \bar{\Psi}\gamma^\mu\Psi := 0. \tag{45}$$

because Ψ fulfills the free Dirac equation. Note that the free electron field is *not* affected by the gauge transformation. We may call $T^\mu_{1/1} = ie : \bar{\Psi}\gamma^\mu\Psi : u$ the 'Q-vertex' of QED. The generalization of (44) to n-th order is

$$[Q, T_n(x_1, ...x_n)] = i\sum_{l=1}^{n} \partial^{x_l}_\mu T^\mu_{n/l}(x_1, ...x_n) = (\text{sum of divergences}) \quad , \tag{46}$$

where $T^\mu_{n/l}$ is again a mathematically well-defined version of the time-ordered product

$$T^\mu_{n/l}(x_1, ..., x_n) " = " T(T_1(x_1)...T^\mu_{1/1}(x_l)...T_1(x_n)). \tag{47}$$

It is relatively easy to see how (46) comes about if one understands the second order case. This example is discussed in detail in Appendix D. We *define* (46) to be the condition of gauge invariance [15].

If we consider for a fixed x_l all terms in T_n with the external field operator $A_\mu(x_l)$

$$T_n(x_1, ...x_n) =: t^\mu_l(x_1, ...x_n)A_\mu(x_l) : +... \tag{48}$$

(the dots represent terms without $A_\mu(x_l)$), then gauge invariance (46) requires

$$\partial^l_\mu[t^\mu_l(x_1, ...x_n)u(x_l)] = t^\mu_l(x_1, ...x_n)\partial_\mu u(x_l) \tag{49}$$

or

$$\partial^l_\mu t^\mu_l(x_1, ...x_n) = 0, \tag{50}$$

i.e. we obtain the Ward-Takahashi identities [16] for QED. The Ward-Takahashi identities express the implications of gauge invariance of QED, which is defined here on the operator level, by C-number identities for Green's distributions.

We have found the following important property of QED: There exists a symmetry transformation generated by the gauge charge Q, which leaves the S-matrix elements invariant, since the gauge transformation only adds divergences in the analytic sense to the S-matrix expansion which vanish after integration over the coordinates $x_1, \ldots x_n$ (see Appendix E).

The S-matrix 'lives' on a Fock space $\mathcal{F}$ which contains physical and unphysical states. E.g., unphysical single photon states can be created by acting on the perturbative vacuum $|\Omega\rangle$ with timelike or longitudinal creation operators:

$$a^0(\vec{k})^\dagger|\Omega\rangle, \quad \sum_{j=1,2,3} k_j a^j(\vec{k})^\dagger|\Omega\rangle, \tag{51}$$

and states which contain ghosts are always unphysical. The physical subspace $\mathcal{F}_{phys}$ is the space of states $|\Phi\rangle$ without timelike and longitudinal photons and ghost, that means

$$a^0(\vec{k})|\Phi\rangle = 0, \quad k_j a^j(\vec{k})|\Phi\rangle = 0, \quad c_{1,2}(\vec{k})|\Phi\rangle = 0 \quad \forall\vec{k}. \tag{52}$$

Note that the definition of the physical space is *not* Lorentz invariant, but the definition (35) is.

The observation that QED is gauge invariant is interesting, but the true importance of gauge invariance is the fact that it allows to prove on a formal level the *unitarity* of the S-matrix on the physical subspace (see the last paper of [15]). Due to the presence of the skew-adjoint operator A^0 or the presence of unphysical longitudinal and timelike photon states, the S-matrix is not unitary on the full Fock space, but it is on $\mathcal{F}_{phys}$. We do not give the algebraic proof here, but we point out that gauge invariance is the basic prerequisite which ensures unitarity. A detailed discussion of this fact can be found in [15, 17, 18]. We have introduced the ghosts only as a formal tool, since they 'blow up' the Fock space unnecessarily, and they do not interact with the electrons and photons. But in QCD, the situation is not so trivial.

6 Quantum chromodynamics

As a further step, we consider QCD without fermions (i.e. quarks) now. The first order coupling of the gluons that we obtain from the classical QCD Lagrangian is given by (see Appendix B)

$$\begin{aligned} T_1^A(x) &= i\frac{g}{2} f_{abc} : A_{\mu a}(x) A_{\nu b}(x) F_c^{\nu\mu}(x) : \\ &= igf_{abc} : A_{\mu a}(x) A_{\nu b}(x) \partial^\nu A_c^\mu(x) :, \quad a, b, c = 1, \dots 8 \end{aligned} \tag{53}$$

where $F_c^{\mu\nu} = \partial^\mu A_c^\nu - \partial^\nu A_c^\mu$ and f_{abc} are the totally antisymmetric structure constants of the gauge group SU(3). The asymptotic free fields satisfy the commutation relations

$$[A_{\mu a}^{(\pm)}(x), A_{\nu b}^{(\mp)}(y)] = i\delta_{ab} g_{\mu\nu} D^{(\mp)}(x-y) \tag{54}$$

and

$$\{u_a^{(\pm)}(x), \tilde{u}_b^{(\mp)}(y)\} = -i\delta_{ab} D^{(\mp)}(x-y), \tag{55}$$

and all other {anti-}commutators vanish. Defining the gauge charge as in (36) by

$$Q := \int d^3x\, \partial_\mu A_a^\mu(x) \overleftrightarrow{\partial_0} u_a(x), \tag{56}$$

where summation over repeated indices is understood, we are led to the following commutators with the fields:

$$[Q, A_a^\mu] = i\partial^\mu u_a, \quad [Q, F_a^{\mu\nu}] = 0, \quad \{Q, u_a\} = 0, \quad \{Q, \tilde{u}_a\} = -i\partial_\mu A_a^\mu \quad . \tag{57}$$

For the commutator of Q with T_1^A we obtain

$$[Q, T_1^A(x)] = igf_{abc} : \{[Q, A_{\mu a}]A_{\nu b}\partial^\nu A_c^\mu + A_{\mu a}[Q, A_{\nu b}]\partial^\nu A_c^\mu + A_{\mu a}A_{\nu b}[Q, \partial^\nu A_c^\mu]\} :$$

$$= -gf_{abc} : \{\partial_\mu u_a A_{\nu b}\partial^\nu A_c^\mu + \partial_\nu(A_{\mu a}u_b\partial^\nu A_c^\mu)\} : . \qquad (58)$$

The last term is a divergence, but the first term spoils gauge invariance and therefore the unitarity of the theory. To restore gauge invariance, we must somehow compensate the first term in (58).

We consider therefore the gluon-ghost coupling term

$$T_1^u(x) = igf_{abc} : A_{\mu a}(x)u_b(x)\partial^\mu \tilde{u}_c(x) :, \qquad (59)$$

For the commutator with Q we get

$$[Q, T_1^u] = ig : [Q, A_{\mu a}]u_b\partial^\mu\tilde{u}_c + A_{\mu a}\{Q, u_b\}\partial^\mu\tilde{u}_c - A_{\mu a}u_b\partial^\mu\{Q, \tilde{u}_c\} :$$

$$= -gf_{abc} : \{\partial_\mu u_a u_b\partial^\mu\tilde{u}_c + A_{\mu a}u_b\partial^\mu\partial^\nu A_{\nu c}\} : . \qquad (60)$$

Note that we can always decompose the commutator in a clever way such that only anticommutators of Q with the ghost fields and commutators with the gauge fields appear. Taking the antisymmetry of f_{abc} and $: u_a u_b := - : u_b u_a :$ into account, we see that the first term is a divergence

$$f_{abc} : \partial_\mu u_a u_b\partial^\mu\tilde{u}_c := \frac{1}{2}f_{abc}\partial_\mu : u_a u_b\partial^\mu\tilde{u}_c : \qquad (61)$$

because $\tilde{u}_c(x)$ fulfills the wave equation. The second term can be written as

$$-gf_{abc} : A_{\mu a}u_b\partial^\mu\partial^\nu A_{\nu c} := -gf_{abc} : \partial^\nu(A_{\mu a}u_b\partial^\mu A_{\nu c}) - \partial^\nu u_b A_{\mu a}\partial^\mu A_{\nu c} : . \qquad (62)$$

Interchanging $a \leftrightarrow b$ and $\mu \leftrightarrow \nu$ in the last term, it becomes equal to the first term in (58).

Hence, the combination

$$T_1(x) = igf_{abc}\{\frac{1}{2} : A_{\mu a}(x)A_{\nu b}(x)F_c^{\nu\mu}(x) : - : A_{\mu a}(x)u_b(x)\partial^\mu\tilde{u}_c(x) :\} \qquad (63)$$

leads to a gauge invariant first order coupling

$$[Q, T_1] = gf_{abc} : -\partial_\nu(A_{\mu a}u_b(\partial^\nu A_c^\mu - \partial^\mu A_c^\nu) + \frac{1}{2}\partial_\nu(u_a u_b\partial^\nu\tilde{u}_c) : . \qquad (64)$$

In contrast to QED, the nonlinear selfinteraction of the unphysical degrees of freedom of the gluons spoils gauge invariance. Coupling ghosts in an appropriate way to the gluons restores the consistency of the theory. Thats why we need ghosts. They interfere destructively with the unphysical gluons and save the unitarity of the theory.

The first order coupling given in (63) is in fact not the most general one, and it is also possible to construct a gauge invariant first order coupling with bosonic ghosts [15]; but gauge invariance then breaks down at second order of perturbation theory. Gauge invariance at order n is again defined by (46). We have discussed gauge invariance only at

first order here. To prove that gauge invariance (and other symmetries of the theory) is not broken by renormalization, is the 'hard problem' of renormalization theory [19, 20, 21].

The 4-gluon term $\sim g^2$ which also appears in the classical Lagrangian is missing in T_1. This term appears as a necessary local normalization term at second order, and its structure is also fixed by perturbative gauge invariance. This stresses the fact that perturbative gauge invariance is strongly related to the formal expansion of the theory in powers of the coupling constant g.

We started this section by presupposing that the coupling of the gluons is already given by the classical Lagrangian. But perturbative gauge invariance is indeed a very restrictive condition: It fixes the interaction to a large extent. We outline here how this fact can be derived. Only a part of the derivation is given here, and it is not assumed that the reader will check the calculations in detail. A full discussion is given in [18], and for further reading we recommend also [22].

We start by describing the interaction of the massless gluons by the most general renormalizable ansatz (i.e. the dimension of the interaction terms is energy^4) with zero ghost number (coupling with non-zero ghost number would affect the theory only on the unphysical sector)

$$\begin{aligned}\tilde{T}_1(x) \quad = ig\Big\{ \quad & \tilde{f}^1_{abc} : A_{\mu a}(x) A_{\nu b}(x) \partial^\nu A^\mu_c(x) : + \\ & \tilde{f}^2_{abc} : A_{\mu a} u_b \partial^\mu \tilde{u}_c : + \\ & \tilde{f}^3_{abc} : A_{\mu a} \partial^\mu u_b \tilde{u}_c : + \\ & \tilde{f}^4_{abc} : A_{\mu a} A^\mu_b \partial_\nu A^\nu_c : + \\ & \tilde{f}^5_{abc} : \partial_\nu A^\nu_a u_b \tilde{u}_c : \Big\} \quad , \quad \tilde{f}^4_{abc} = \tilde{f}^4_{bac},\end{aligned}$$

where the $\tilde{f}$'s are arbitrary real constants. This first order coupling term is then antisymmetric with respect to the conjugation K defined in Appendix C.

Adding divergence terms to $\tilde{T}_1$ will not change the physics of the theory, since divergences get integrated out in (33), and we assume that this property is not destroyed by the renormalization of the theory. Furthermore, the (anti-)commutator of Q with a divergence is also a divergence. We can therefore always calculate modulo divergences in the following. We modify $\tilde{T}_1$ by adding $-\frac{ig}{4}(\tilde{f}^1_{abc} + \tilde{f}^1_{cba})\partial_\nu : A_{\mu a} A_{\nu b} A^\mu_c :$ and $-ig\tilde{f}^3_{abc}\partial_\mu : A^\mu_a u_b \tilde{u}_c :$ and arrive at a more compact, equivalent first order coupling

$$\begin{aligned}T_1 \quad = ig\Big\{ \quad & f^1_{abc} : A_{\mu a} A_{\nu b} \partial^\nu A^\mu_c : + \\ & f^2_{abc} : A_{\mu a} u_b \partial^\mu \tilde{u}_c : + \\ & f^4_{abc} : A_{\mu a} A^\mu_b \partial_\nu A^\nu_c : + \\ & f^5_{abc} : \partial_\nu A^\nu_a u_b \tilde{u}_c : \Big\},\end{aligned} \tag{65}$$

where

$$f^1_{abc} = -f^1_{cba} \quad f^4_{abc} = f^4_{bac}. \tag{66}$$

The gauge variation of T_1 is

$$
\begin{aligned}
[Q, T_1] \quad = -g f^1_{abc} \quad & \Big\{ : \partial_\mu u_a A_{\nu b} \partial^\nu A^\mu_c + \\
& A_{\mu a} \partial_\nu u_b \partial^\nu A^\mu_c + \\
& A_{\mu a} A_{\nu b} \partial^\nu \partial^\mu u_c : \Big\} + \\
-g f^2_{abc} \quad & \Big\{ : \partial_\mu u_a u_b \partial^\mu \tilde{u}_c + \\
& A_{\mu a} u_b \partial^\mu \partial_\nu A^\nu_c : \Big\} + \\
-2g f^4_{abc} \quad & : \partial_\mu u_a A^\mu_b \partial_\nu A^\nu_c : + \\
-g f^5_{abc} \quad & : \partial_\nu A^\nu_a u_b \partial_\mu A^\mu_c : . \qquad (67)
\end{aligned}
$$

Gauge invariance requires that the gauge variation be a divergence, hence we make again a general ansatz

$$
\begin{aligned}
[Q, T_1] \quad = g \partial_\mu \{ \quad & g^1_{abc} : \partial^\mu u_a A_{\nu b} A^\nu_c : + \\
& g^2_{abc} : u_a A^\mu_b \partial_\nu A^\nu_c : + \\
& g^3_{abc} : \partial^\mu u_a u_b \tilde{u}_c : + \\
& g^4_{abc} : u_a u_b \partial^\mu \tilde{u}_c : + \\
& g^5_{abc} : \partial_\nu u_a A^\nu_b A^\mu_c : + \\
& g^6_{abc} : u_a \partial^\mu A^\nu_b A_{\nu c} : + \\
& g^7_{abc} : u_a \partial^\nu A^\mu_b A_{\nu c} : \}, \qquad (68)
\end{aligned}
$$

where $g^1_{abc} = g^1_{acb}, g^4_{abc} = -g^4_{bac}$. Comparing the terms in (67) and (68), we obtain a set of constraints for the coupling coefficients:

$$
\begin{aligned}
-f^1_{cab} &= 2g^1_{abc} + g^6_{abc} \quad \text{from} \quad && : \partial^\mu u \partial_\mu A_\nu A^\nu : && (69) \\
g^2_{abc} + g^5_{abc} &= -2f^4_{abc} && : \partial^\mu u A_\mu \partial_\nu A^\nu : && (70) \\
g^2_{abc} + g^2_{acb} &= -f^5_{bac} - f^5_{cab} && : u \partial_\nu A^\nu \partial_\mu A^\mu : && (71) \\
-f^2_{abc} &= g^3_{abc} + 2g^4_{abc} && : \partial_\mu u u \partial^\mu \tilde{u} : && (72) \\
g^3_{abc} &= g^3_{bac} && : \partial_\mu u \partial^\mu u \tilde{u} : && (73) \\
-f^1_{cba} - f^1_{bca} &= g^5_{abc} + g^5_{acb} && : \partial_\mu \partial_\nu u A^\nu A^\mu : && (74) \\
-f^1_{acb} &= g^5_{abc} + g^7_{abc} && : \partial_\mu u \partial_\nu A^\mu A^\nu : && (75) \\
g^6_{abc} &= -g^6_{acb} && : u \partial_\mu A_\nu \partial^\mu A^\nu : && (76) \\
-f^2_{cab} &= g^2_{acb} + g^7_{abc} && : u \partial_\mu \partial_\nu A^\mu A^\nu : && (77) \\
g^7_{abc} &= -g^7_{acb} && : u \partial_\nu A^\mu \partial^\mu A^\nu : . && (78)
\end{aligned}
$$

From (74), (75), and (78) we readily derive

$$
f^1_{abc} + f^1_{acb} - f^1_{bca} - f^1_{cba} = 0. \qquad (79)
$$

Combining this with (66), we obtain the first important result that f^1_{abc} is *totally antisymmetric*. g^1_{abc} is symmetric in b and c; from (76) and (69) we conclude $g^1_{abc} = 0$ and $g^6_{abc} = -f^1_{abc}$.

We did not yet fully exploit gauge invariance. Taking into account additionally that the gauge charge is nilpotent $Q^2 = 0$ and consequently

$$\{Q, [Q, T_1]\} = 0 \tag{80}$$

provides further relations which fix the interaction to

$$T_1 = T_1^{YM} + T_1^D, \tag{81}$$

$$T_1^{YM} = igf^1_{abc}\{: \frac{1}{2}A_{\mu a}A_{\nu b}F_c^{\nu\mu} : - : A_{\mu a}u_b\partial^\mu\tilde{u}_c :\}, \tag{82}$$

$$T_1^D = igf^5_{abc} : \partial_\nu A_a^\nu u_b\tilde{u}_c :, \tag{83}$$

where 'YM' stands for 'Yang-Mills' and 'D' for 'Deformation', and $f^5_{abc} = -f^5_{cba}$.

We stop our analysis here for the sake of brevity. A detailed analysis of gauge invariance *at second order* shows [18] that the coupling coefficients f_{abc} must fulfill the Jacobi identity

$$f_{abc}f_{dec} + f_{adc}f_{ebc} + f_{aec}f_{bdc} = 0, \tag{84}$$

i.e. we are lead to the nice result that the f_{abc} are structure constants of a Lie group. Note that the Lie structure of gauge theories is usually presupposed, but here it follows from basic symmetries of quantum field theory, namely quantum gauge invariance (which is related to unitarity) and Lorentz covariance.

Finally, we are left with the term

$$T_1^D = igf^5_{abc} : \partial_\nu A_a^\nu u_b\tilde{u}_c : . \tag{85}$$

One can show that this interaction term which contains only unphysical fields does not contribute to the S-matrix on the physical sector.

The reason why gauge invariance at higher orders of the perturbation expansion is a nontrivial task can be understood qualitatively on a basic level. Commutators of free fields can be expressed by (derivatives of) Pauli-Jordan distributions, e.g. we have

$$\{\partial_\nu A^\nu(x), \partial_\mu A^\mu(y)] = ig^{\nu\mu}\partial_\nu^x\partial_\mu^y D(x-y)$$

$$= -i\Box_x D(x-y) = 0. \tag{86}$$

But through the timeordering process, Feynman type propagators $\sim D_F(x-y)$ appear in the amplitudes instead of Pauli-Jordan distributions, which fulfill the inhomogeneous wave equation

$$\Box D_F(z) = -\delta^{(4)}(z), \tag{87}$$

i.e. derivatives acting on Feynman propagators may generate 'anomalous' local terms, because time ordering does not commute with analytic derivation. Therefore, maintaining gauge invariance may imply restrictions to the coupling structure also at higher orders. If it is not possible to fix the theory such that gauge invariance can be maintained, the theory is called *nonrenormalizable*. The result obtained by the analysis above can be considered as an inversion of 't Hooft's famous result that Yang-Mills theories are renormalizable [23].

In this section, we focused on the purely gluonic part of QCD. A full discussion would include also the coupling of gluons to quarks, but all important features are contained in the present discussion.

7 Yang-Mills theories with massive gauge fields

In this section we give a short discussion of a general Yang-Mills theory with massive gauge bosons. Free massive spin 1 fields with color index a satisfy the Proca equation

$$\Box A^\mu_a - \partial^\mu(\partial_\nu A^\nu_a) + m_a^2 A^\mu_a = 0. \tag{88}$$

Taking the divergence of (88) shows that the field automatically satisfies the Lorentz condition $\partial_\mu A^\mu_a = 0$. Again we may quantize the fields as four independent scalar fields which fulfill the wave equation

$$(\Box + m_a^2) A^\mu_a = 0, \tag{89}$$

and the Lorentz condition shall be enforced on the Fock space like in the massless case by selecting a physical subspace which contains the three allowed polarizations. The definition of the gauge charge via (36) does no longer lead to a nilpotent gauge charge due to the mass of the gauge field. This can be restored by introducing scalar bosonic fields Φ_a with the same mass as the corresponding gauge field

$$(\Box + m_a^2)\Phi_a = 0. \tag{90}$$

These fields are called Higgs ghosts or would-be Goldstone bosons, and they are unphysical. We demand that the free fields then satisfy the commutation relations

$$[A^\mu_a(x), A^\nu_b(y)] = i\delta_{ab} g^{\mu\nu} D_{m_a}(x-y), \tag{91}$$

$$[u_a(x), \tilde{u}_b(y)] = -i\delta_{ab} D_{m_a}(x-y), \tag{92}$$

$$[\Phi_a(x), \Phi_b(y)] = -i\delta_{ab} D_{m_a}(x-y), \tag{93}$$

where $D_{m_a}(x-y)$ are the usual Pauli-Jordan distributions for massive fields

$$D_m(x) = \frac{i}{(2\pi)^3} \int d^4k \, \delta(k^2 - m^2) \mathrm{sgn}(k^0) e^{-ikx}. \tag{94}$$

Then we can define a nilpotent gauge charge by

$$Q := \int d^3x \, (\partial_\mu A^\mu_a(x) + m_a \Phi_a(x)) \overleftrightarrow{\partial_0} \, u(x) \tag{95}$$

Gauge variations are then obtained from the commutators

$$[Q, A^\mu_a] = i\partial^\mu u_a, \quad \{Q, u_a\} = 0,$$

$$\{Q, \tilde{u}_a\} = -i(\partial_\mu A^\mu_a + m_a \Phi_a), \quad [Q, \Phi_a] = i m_a u_a. \tag{96}$$

The coupling (63) which we obtained in the massless case

$$T_1(x) = igf_{abc}\{\frac{1}{2} : A_{\mu a}(x) A_{\nu b}(x) F^{\nu\mu}_c(x) : - : A_{\mu a}(x) u_b(x) \partial^\mu \tilde{u}_c(x) :\} \tag{97}$$

is not gauge invariant anymore due to the additional massterms in the gauge variations (96). But gauge invariance can be restored at least at first order if we add to T_1 a Higgs ghost coupling

$$T_1^\Phi = i\frac{g}{2} f_{abc}\Big(\frac{m_b^2+m_c^2-m_a^2}{m_b m_c} :A_{\mu a}\Phi_b\partial^\mu\Phi_c:$$

$$+\frac{m_a^2+m_c^2-m_b^2}{m_a} :\Phi_a u_b\tilde{u}_c: +\frac{m_b^2-m_a^2}{m_c} :A_{\mu a}A_b^\mu\Phi_c:\Big)\,, \tag{98}$$

for the case where all masses are nonvanishing. But this is not the full story, because gauge invariance can be shown to break down again at second order under certain circumstances. As a consequence, a complete analysis of the situation is necessary also at higher orders. In order to illustrate the implications of gauge invariance, we present here the result of such an analysis for the case where we have three different colors, and we assume that at least one of them is massive (because otherwise we would be dealing with a massless SU(2) Yang-Mills theory).

It is found that there exist two *minimal scenarios*, which are gauge invariant also at second order, whereas the coupling f_{abc} is equal to the completely antisymmetric tensor ε_{abc} for both cases:

One possibility is given by the case where two gauge boson are equally massive, and the third one is massless. The correct coupling structure can be obtained from (98) by performing a limit $m_3 \to 0$ for $m_1 = m_2$. The corresponding massless Higgs ghost does not appear anymore in the gauge charge Q, and becomes a massless physical field. The second possibility is given by the case where all three gauge bosons are massive, and all masses are then equal. But it turns also out that it is necessary to introduce additional fields in order to save gauge invariance. Indeed, it is already sufficient to add one physical scalar to the theory (i.e. this is the *minimal* solution): A Higgs boson. Gauge invariance fixes the structure of the coupling completely, but the mass of the Higgs remains a free parameter in the theory. The two scenarios correspond to the classical picture of the two possible types of spontaneous symmetry breaking for a SU(2) Yang-Mills theory: In the first case, one couples a real SO(3) field triplet to the gauge fields. Two degrees of freedom of this triplet are 'eaten up' by the gauge bosons, which become massive. In the second case, one couples a complex doublet (equivalent to four real fields) to the gauge bosons. Three degrees of freedom get absorbed by the gauge fields, which become massive, and one degree of freedom remains as a physical Higgs field. These two mechanisms are well described in many textbooks (see e.g. [24], chapter 8.3). A detailed discussion in the present approach of theories involving massive gauge fields like the Z- and $W^\pm$-bosons can be found in [25, 26, 27, 28, 29]. The interesting point of the free field approach is given by the fact that it reverses the chain of arguments usually presented in the literature. Fields do not become massive due to the coupling to a Higgs field, but the existence of a Higgs field becomes necessary due to the mass of the gauge bosons and as a consequence of perturbative gauge invariance.

8 Quantum gravity (as massless spin 2 gauge theory)

For the following discussion, we should always keep in mind that quantum gravity is a nonrenormalizable theory, and not really understood to all orders in perturbation theory.

But at least it is possible to discuss gauge invariance of the lowest orders [30, 31].

The general theory of relativity can be derived from the Einstein-Hilbert Lagrangian

$$\mathcal{L}_{EH} = -\frac{2}{\kappa^2}\sqrt{-g}R \tag{99}$$

where $R = g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar, g is the determinant of the metric tensor $g^{\mu\nu}$ and $\kappa^2 = 32\pi G$, where G is Newton's gravitational constant. It is convenient to introduce Goldberg variables [32]

$$\tilde{g}^{\mu\nu} = \sqrt{-g}g^{\mu\nu}. \tag{100}$$

Since we are working on a perturbative level, we expand the inverse metric $\tilde{g}^{\mu\nu}$ and assume that we have an asymptotically flat geometry

$$\tilde{g}^{\mu\nu} = \eta^{\mu\nu} + \kappa h^{\mu\nu}. \tag{101}$$

Here $\eta^{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$ is the metric of Minkowski spacetime. We used the more common symbol $g^{\mu\nu}$ for the flat space metric in the previous sections, but use $\eta^{\mu\nu}$ in the following in order to point out the difference between η and g more clearly. All tensor indices will be raised and lowered with $\eta^{\mu\nu}$. The quantity $h^{\mu\nu}$ is a symmetric second rank tensor field, which describes gravitons after quantization. For notational convenience, we do not care if the indices are up or down, since the summation convention always makes clear what is meant. Formally (99) can be expanded as an infinite power series in κ

$$\mathcal{L}_{EH} = \sum_{j=0}^{\infty} \kappa^j \mathcal{L}_{EH}^{(j)}. \tag{102}$$

The lowest order term $\mathcal{L}_{EH}^{(0)}$ is quadratic in $h^{\mu\nu}(x)$ and defines the wave equation of the free asymptotic fields. The linearized Euler-Lagrange equations of motion for $h^{\mu\nu}(x)$ are

$$\Box h^{\mu\nu}(x) - \frac{1}{2}\eta^{\mu\nu}\Box h(x) - h^{\mu\rho,\nu}_{,\rho}(x) - h^{\nu\rho,\mu}_{,\rho}(x) = 0 \tag{103}$$

where $h(x) = h^{\mu}_{\mu}(x)$, and the commas denote partial derivatives with respect to the corresponding index. This equation is invariant under a classical gauge transformations of the form

$$h^{\mu\nu} \longrightarrow h^{\mu\nu} + u^{\mu,\nu} + u^{\nu,\mu} - \eta^{\mu\nu}u^{\rho}_{,\rho} \tag{104}$$

where u^μ is a vector field which satisfies the wave equation

$$\Box u^\mu(x) = 0. \tag{105}$$

As a consequence, quantum gauge invariance for gravity will be formulated by the help of (fermionic) vector ghosts. This gauge transformation corresponds to the general covariance in its linearized form of the metric tensor $g_{\mu\nu}(x)$. The corresponding gauge condition, compatible with (104) is the Hilbert-gauge, which obviously plays a similar role as the Lorentz gauge condition for spin 1 fields

$$h^{\mu\nu}_{,\mu} = 0. \tag{106}$$

Then the dynamical equation for the graviton field $h^{\mu\nu}$ reduces to the wave equation

$$\Box h^{\mu\nu}(x) = 0. \tag{107}$$

The first order term with respect to κ, $\mathcal{L}^{(1)}_{EH}$ gives the trilinear self-coupling of the gravitons

$$\mathcal{L}^{(1)}_{EH} = \frac{\kappa}{2}\, h^{\rho\sigma}\Big(h^{\alpha\beta}_{,\rho} h^{\alpha\beta}_{,\sigma} - \frac{1}{2}\, h_{,\rho} h_{,\sigma} + 2\, h^{\alpha\rho}_{,\beta} h^{\beta\sigma}_{,\alpha} + h_{,\alpha} h^{\rho\sigma}_{,\alpha} - 2\, h^{\alpha\rho}_{,\beta}\, h^{\alpha\sigma}_{,\beta}\Big). \tag{108}$$

There exists many alternative derivations of this result (108), starting from massless tensor fields and requiring consistency with gauge invariance in some sense. A similar point of view is the one of Ogievetsky and Polubarinov [33]. They analyzed spin 2 theories by working with a generalized Hilbert-gauge condition to exclude the spin one part from the outset. They imposed an invariance under infinitesimal gauge transformations of the form

$$\delta h^{\mu\nu} = \partial^\mu u^\nu + \partial^\nu u^\mu + \eta^{\mu\nu}\partial_\alpha u^\alpha \tag{109}$$

and obtained Einstein's theory at the end. Instead Wyss [34] considered the coupling to matter. Then the self-coupling of the tensor-field (108) is necessary for consistency. Wald [35] derived a divergence identity which is equivalent to an infinitesimal gauge invariance of the theory. Einstein's theory is the only non-trivial solution of this identity. In quantum theory the problem was studied by Boulware and Deser [36]. These authors require Ward identities associated with the graviton propagator to implement gauge invariance. All authors get Einstein's theory as the unique classical limit if the theory is purely spin two (without a spin one admixture).

The free asymptotic field $h^{\mu\nu}$ is a symmetric tensor field of rank two and u^μ and $\tilde{u}^\nu$ are ghost and anti-ghost fields on the background of Minkowski spacetime in the following. A symmetric tensor field has ten degrees of freedom, which are more than the five independent components of a massive spin 2 field or the two degrees of freedom of the massless graviton field. The additional degrees of freedom can be eliminated on the classical level by imposing further conditions, but we will now quantize the graviton field and introduce the gauge charge operator Q. The physical states in the full Fock space which contains eight unphysical polarizations of the graviton and ghosts can then be characterized directly by the help of the gauge charge as in the case of a spin 1 field (see Appendix C, and [37, 38, 39]).

The tensor field $h^{\mu\nu}(x)$ can be quantized as a massless field as follows

$$\big[h^{\alpha\beta}(x), h^{\mu\nu}(y)\big] = -ib^{\alpha\beta\mu\nu} D_0(x-y) \tag{110}$$

where $D_0(x-y)$ is again the massless Pauli-Jordan distribution and the tensor $b^{\alpha\beta\mu\nu}$ is constructed from the Minkowski metric $\eta^{\mu\nu}$ via

$$b^{\alpha\beta\mu\nu} = \frac{1}{2}\big(\eta^{\alpha\mu}\eta^{\beta\nu} + \eta^{\alpha\nu}\eta^{\beta\mu} - \eta^{\alpha\beta}\eta^{\mu\nu}\big). \tag{111}$$

An explicit representation of the field $h^{\mu\nu}(x)$ on the Fock space is given by

$$h^{\alpha\beta}(x) = (2\pi)^{-3/2}\int \frac{d^3k}{\sqrt{2\omega}}\Big(a^{\alpha\beta}(\vec{k})e^{-ikx} + a^{\alpha\beta}(\vec{k})^\dagger e^{+ikx}\Big). \tag{112}$$

Here we have $\omega = |\vec{k}|$ as usual and $a^{\alpha\beta}(\vec{k})$, $a^{\alpha\beta}(\vec{k})^\dagger$ are annihilation and creation operators on a bosonic Fock-space. One finds that they have the following commutation relations

$$\left[a^{\alpha\beta}(\vec{k}), a^{\mu\nu}(\vec{k}')^\dagger\right] = b^{\alpha\beta\mu\nu}\delta^{(3)}(\vec{k}-\vec{k}'), \tag{113}$$

and all other commutators vanish.

In analogy to the spin 1 theories like QCD we may introduce a gauge charge operator by

$$Q := \int\limits_{x^0=const.} d^3x\, h^{\alpha\beta}(x)_{,\beta}\overset{\leftrightarrow}{\partial}_0 u_\alpha(x). \tag{114}$$

For the construction of the physical subspace and in order to prove the unitarity of the S-matrix we must have a nilpotent operator Q. Therefore we have to quantize the ghost fields with anticommutators according to

$$\left\{u^\mu(x), \tilde{u}^\nu(y)\right\} = i\eta^{\mu\nu}D_0(x-y). \tag{115}$$

All asymptotic fields fulfill the wave equation

$$\Box\, h^{\mu\nu}(x) = \Box\, u^\alpha(x) = \Box\, \tilde{u}^\beta(x) = 0. \tag{116}$$

We obtain the following commutators of the fundamental fields:

$$[Q, h^{\mu\nu}] = -\frac{i}{2}\left(u^\mu_{,\nu} + u^\nu_{,\mu} - \eta^{\mu\nu}u^\alpha_{,\alpha}\right), \tag{117}$$

$$[Q, h] = iu^\mu_{,\mu}, \tag{118}$$

$$\{Q, \tilde{u}^\mu\} = ih^{\mu\nu}_{,\nu}, \tag{119}$$

$$\{Q, u^\mu\} = 0. \tag{120}$$

From (117) we immediately get

$$[Q, h^{\mu\nu}_{,\mu}] = 0. \tag{121}$$

The result (117) agrees with the infinitesimal gauge transformations of the Goldberg variables, so that our quantization and choice of Q corresponds to the classical framework.

Again, first order gauge invariance means that $[Q, T_1]$ is a divergence in the sense of vector analysis, i.e.

$$[Q, T_1(x)] = i\partial_\mu T^\mu_{1/1}(x), \tag{122}$$

and definition of the n-th order gauge invariance then reads

$$[Q, T_n] = \left[Q, T_n\right] = i\sum_{l=1}^{n}\frac{\partial}{\partial x^\mu_l}T^\mu_{n/l}(x_1,\ldots,x_l,\ldots,x_n). \tag{123}$$

Like in the case of QCD, we are forced to add a ghost part to the self-coupling term of the gravitons. The first order graviton coupling according to (108) is

$$\tilde{T}^h_1 = \frac{i\kappa}{2}\,h^{\rho\sigma}\Big(h^{\alpha\beta}_{,\rho}h^{\alpha\beta}_{,\sigma} - \frac{1}{2}\,h_{,\rho}h_{,\sigma} + 2\,h^{\alpha\rho}_{,\beta}h^{\beta\sigma}_{,\alpha} + h_{,\alpha}h^{\rho\sigma}_{,\alpha} - 2\,h^{\alpha\rho}_{,\beta}h^{\alpha\sigma}_{,\beta}\Big). \tag{124}$$

By adding a physically irrelevant divergence to $\tilde{T}_1^h$, we obtain a more compact expression for the graviton interaction

$$T_1^h(x) = i\kappa : \Big(\frac{1}{2}h^{\mu\nu}h^{\alpha\beta}_{,\mu}h^{\mu\nu}h^{\alpha\beta}_{,\nu} + h^{\mu\nu}h^{\nu\alpha}_{,\beta}h^{\mu\beta}_{,\alpha} - \frac{1}{4}h^{\mu\nu}h_{,\nu}h_{,\nu}\Big) : . \tag{125}$$

The ghost coupling turns out to be the one first suggested by Kugo and Ojima [40], namely

$$T_1^u = i\kappa : \tilde{u}^\nu_{,\mu}\Big(h^{\mu\nu}_{,\rho}u^\rho - h^{\nu\rho}u^\mu_{,\rho} - h^{\mu\rho}u^\nu_{,\rho} + h^{\mu\nu}u^\rho_{,\rho}\Big) : . \tag{126}$$

It is a nice detail that the four graviton vertex which follows from the second order term in (102) is also proliferated by gauge invariance at second order. It is therefore quite probable that also all higher vertex couplings which appear in (102) are proliferated by quantum gauge invariance. It is also possible to derive the couplings (124),(126) from perturbative gauge invariance similarly as it was done in sect. 6 for Yang-Mills theories. For this and further reading we refer to the recent monograph [22].

We finally point out that it would be interesting to analyze the Higgs mechanism for massive gravity [41] in the framework presented above, also in connection with the dark matter problem. A first step in this direction is presented in [42].

9 Appendices

Appendix A: Commutator of the gauge charge Q with gauge fields

First we derive the distributional identity (16). The Pauli-Jordan distribution has an integral representation

$$D(x) = \frac{i}{(2\pi)^3}\int d^4k\,\delta(k^2)\mathrm{sgn}(k_0)e^{-ikx} \tag{127}$$

Using the identity

$$\delta(k^2) = \delta(k_0^2 - \vec{k}^{\,2}) = \frac{1}{2|k^0|}\Big(\delta(k^0 - |\vec{k}|) + \delta(k^0 + |\vec{k}|)\Big), \tag{128}$$

we obtain (mind the $\mathrm{sgn}(k^0)$)

$$\partial_0 D(x) = \frac{i}{(2\pi)^3}\int\frac{d^4k}{2|k^0|}\Big(\delta(k^0 - |\vec{k}|) - \delta(k^0 + |\vec{k}|\Big)(-ik^0)e^{-ikx}$$

$$\frac{1}{2(2\pi)^3}\int d^3k\,\Big(e^{-i(|\vec{k}|x^0 - \vec{k}\vec{x})} + e^{-i(-|\vec{k}|x^0 - \vec{k}\vec{x})}\Big). \tag{129}$$

Restricting this result to $x^0 = 0$, we get the desired result (16)

$$\partial_0 D(x)|_{x^0=0} = (2\pi)^{-3}\int d^3k\,e^{+i\vec{k}\vec{x}} = \delta^{(3)}(\vec{x}). \tag{130}$$

In a completely analogous way, one derives

$$\partial_0^2 D(x)|_{x^0=0} = 0, \quad \vec{\nabla}D(x)|_{x^0=0} = 0. \tag{131}$$

Note that we always consider the well-defined differentiated distribution first, which gets then restricted to a subset of its support.

As an example, we investigate now the commutator $[Q, A_\mu(y)]$, and we omit the trivial color index. The commutator is given explicitly by

$$[Q, A_\mu(y)] = [\int\limits_{x^0=y^0} d^3x\, \partial_\nu A^\nu(x) \overleftrightarrow{\partial_0^x} u(x), A_\mu(y)] = i \int\limits_{x^0=y^0} d^3x\, \partial_\mu^x D(x-y) \overleftrightarrow{\partial_0^x} u(x). \quad (132)$$

Here, we made use of the freedom to choose any constant value for x^0, and therefore we set $x^0 = y^0$, such that we have $x^0 - y^0 = 0$ and we can apply (16, 131) in the sequel.

For $\mu = 0$, we have

$$[Q, A_0(y)] = i \int\limits_{x^0=y^0} d^3x\, \partial_0^x D(x-y) \overleftrightarrow{\partial_0^x} u(x)$$

$$= i \int\limits_{x^0=y^0} d^3x\, \delta^{(3)}(\vec{x}-\vec{y}) \partial_0^x u(x) = i\partial_0 u(y), \quad (133)$$

where we have used that the double timelike derivative of D vanishes on the integration domain according to (131). The result for the commutator of Q with the spacelike components of A is also obtained by using (16, 131) and by shifting the gradient acting of the Pauli-Jordan distribution by partial integration on the ghost field.

Appendix B: The Becchi-Rouet-Stora-Tyutin transformation and its free field version

The gluon vector potential can be represented by the traceless hermitian 3×3 standard Gell-Mann matrices λ^a, $a = 1, ...8$

$$A_\mu = \sum_{a=1}^{8} A_\mu^a \frac{\lambda^a}{2} =: A_\mu^a \frac{\lambda^a}{2}. \quad (134)$$

The λ's satisfy the commutation and normalization relations

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = i f_{abc} \frac{\lambda^c}{2}, \quad \mathrm{tr}\,(\lambda^a \lambda^b) = 2\delta_{ab}, \quad (135)$$

and the numerical values of the structure constants $f_{abc} = -f_{bac} = -f_{acb}$ can be found in numerous QCD textbooks. Since we are working with a fixed matrix representation, we do not care whether the color indices are upper or lower indices.

The natural generalization of the QED Lagrangian to the Lagrangian of purely gluonic QCD is

$$\mathcal{L}_{gluon} = -\frac{1}{2} \mathrm{tr}\, G_{\mu\nu} G^{\mu\nu} = -\frac{1}{4} G^a_{\mu\nu} G_a^{\mu\nu}, \quad (136)$$

with

$$G_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \quad (137)$$

or, using the first relation of (135)

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f_{abc} A^b_\mu A^c_\nu. \tag{138}$$

It is an important detail that we are working with *interacting* classical fields here, therefore the field strength tensor contains a term proportional to the coupling constant in contrast to the free field tensor $F^{free}_{\mu\nu} = \partial_\mu A^{free}_\nu - \partial_\nu A^{free}_\mu$ used in this paper. $\mathcal{L}_{gluon}$ is invariant under classical local gauge transformations

$$A_\mu(x) \to U(x) A_\mu(x) U^{-1}(x) + \frac{i}{g} U(x) \partial_\mu U^{-1}(x), \tag{139}$$

where $U(x) \in SU(3)$.

We extract now the first order gluon coupling from the Lagrangian. The Lagrangian

$$\mathcal{L}_{gluon} = -\frac{1}{4}[\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f_{abc} A^b_\mu A^c_\nu][\partial^\mu A^\nu_a - \partial^\nu A^\mu_a + g f_{ab'c'} A^\mu_{b'} A^\nu_{c'}], \tag{140}$$

contains obviously the free field part (this terminology is not really correct, since we are dealing with interacting fields here)

$$\mathcal{L}^{free}_{gluon} = -\frac{1}{4}[\partial_\mu A^a_\nu - \partial_\nu A^a_\mu][\partial^\mu A^\nu_a - \partial^\nu A^\mu_a] \tag{141}$$

and the first order interaction part is given by

$$\mathcal{L}^{int}_{gluon} = -\frac{1}{4}[\partial_\mu A^a_\nu - \partial_\nu A^a_\mu][g f_{ab'c'} A^\mu_{b'} A^\nu_{c'}] - \frac{1}{4}[g f_{abc} A^b_\mu A^c_\nu][\partial^\mu A^\nu_a - \partial^\nu A^\mu_a]$$

$$= -\frac{g}{2} f_{abc} A^b_\mu A^c_\nu [\partial^\mu A^\nu_a - \partial^\nu A^\mu_a] = -\frac{g}{2} f_{abc} A^a_\mu A^b_\nu [\partial^\mu A^\nu_c - \partial^\nu A^\mu_c]$$

$$= g f_{abc} A^a_\mu A^b_\nu \partial^\nu A^\mu_c. \tag{142}$$

From this term follows the first order interaction ansatz $T^A_1 = i : \mathcal{L}^{int}_{gluon} :$ (53).

Since we are working in Feynman gauge, we add the corresponding gauge fixing term $\mathcal{L}_{gf}$ to the Lagrangian. Additionally, we add a ghost term which leads to the ghost interaction (59). The total Lagrangian then reads

$$\mathcal{L}_{QCD} = \mathcal{L}_{gluon} + \mathcal{L}_{gf} + \mathcal{L}_{ghost}$$

$$= \mathcal{L}_{gluon} - \frac{1}{2}(\partial_\mu A^\mu_a)^2 + \partial^\mu \tilde{u}(\partial_\mu u_a - g f_{abc} u_b A_{\mu c}). \tag{143}$$

The classical ghosts are anticommuting Grassmann numbers, i.e. $u^2 = \tilde{u}^2 = 0, u\tilde{u} = -\tilde{u}u$.

The BRST transformation is defined by

$$\delta A^a_\mu = i\lambda(\partial_\mu u_a - g f_{abc} u_b A_{\mu c}), \tag{144}$$

$$\delta \tilde{u}_a = -i\lambda \partial_\mu A^\mu_a, \tag{145}$$

$$\delta u_a = \frac{g}{2}\lambda f_{abc} u_b u_c, \tag{146}$$

where λ is a space-time independent anticommuting Grassmann variable. The special property of the BRST transformation is the fact that the actions

$$S_{gluon} = \int d^4x \, \mathcal{L}_{gluon}, \quad S_{gf} + S_{ghost} = \int d^4x \, (\mathcal{L}_{gf} + \mathcal{L}_{ghost}) \tag{147}$$

and $S_{total} = S_{gluon} + S_{gf} + S_{ghost}$ are all invariant under the transformation:

$$\delta S_{gluon} = 0, \quad \delta(S_{gf} + S_{ghost}) = 0. \tag{148}$$

The similarity of free quantum gauge transformation introduced in this paper to the BRST transformation is obvious. One important difference is the absence of interaction terms $\sim g$. Furthermore, the free quantum gauge transformation is a transformation of free quantum fields, whereas the BRST transformation is a transformation of classical fields, which enter in path integrals when the theory is quantized. Finally, the free gauge transformation leaves the T_n's invariant up to divergences, whereas the BRST transformation is a symmetry of the full QCD Lagrangian. How the two symmetries are intertwined perturbatively is explained in [20]. A more rigorous axiomatic approach is discussed in [43, 19].

Appendix C: Some technical remarks concerning the gauge charge Q

Using the Leibnitz rule for graded algebras gives for the gauge charge for massless spin 1 fields

$$Q^2 = \frac{1}{2}\{Q, Q\} = \frac{1}{2} \int\limits_{x_0 = const.} d^3x \partial_\nu A_a^\nu(x) \{\overset{\leftrightarrow}{\partial}_{x_0} u_a(x), Q\}$$

$$-\frac{1}{2} \int\limits_{x_0 = const.} d^3x [\partial_\nu A_a^\nu(x), Q] \overset{\leftrightarrow}{\partial}_{x_0} u_a(x) = 0 \quad , \tag{149}$$

i.e. Q is nilpotent. This basic property of Q, which holds also in the massive case, and the so-called Krein structure on the Fock-Hilbert space [44, 45] allows to prove unitarity of the S-matrix on the physical Hilbert space $\mathcal{F}_{phys}$, which is a subspace of the Fock-Hilbert space $\mathcal{F}$ containing also the unphysical ghosts and unphysical degrees of freedom of the vector field [46].

The physical Fock space can be expressed by the kernel and the range of Q [46, 44], namely

$$F_{phys} = ker\, Q / ran\, Q = ker\{Q, Q^\dagger\} \quad . \tag{150}$$

Again, this characterization of the physical space holds also for the massive spin 1 case, and for the gauge charge (114) used for gravity. The Krein structure is defined by introducing a conjugation K

$$a_0(\vec{k})^K = -a_0(\vec{k})^\dagger, \quad a_j(\vec{k})^K = a_j(\vec{k})^\dagger, \tag{151}$$

so that $A_\mu^K = A_\mu$, and on the ghost sector

$$c_2(\vec{k})^K = c_1(\vec{k})^\dagger, \quad c_1(\vec{k})^K = c_2(\vec{k})^\dagger, \tag{152}$$

so that $u^K = u$ is K-selfadjoint and $\tilde{u}^K = -\tilde{u}$. Then Q is densely defined on the Fock-Hilbert space and becomes K-symmetric $Q \subset Q^K$. Roughly speaking, the K-conjugation is the natural generalization of the usual hermitian conjugation to the full (unphysical) Fock space.

A calculation shows that the anticommutator in (150) is essentially the number operator for unphysical particles

$$\{Q^\dagger, Q\} = 2\int d^3k\, \vec{k}^2 \big[b_1^\dagger(\vec{k})b_1(\vec{k}) + b_2^\dagger(\vec{k})b_2(\vec{k}) + c_1^\dagger(\vec{k})c_1(\vec{k}) + c_2^\dagger(\vec{k})c_2(\vec{k})\big], \tag{153}$$

with

$$b_{1,2} = (a_\parallel \pm a_0)/\sqrt{2} \quad , \quad a_\parallel = k_j a_j/|\vec{k}|, \tag{154}$$

which implies (150).

The nilpotency of Q allows for standard homological notions [46]: Consider the field algebra $\mathcal{F}$ consisting of the polynomials in the (smeared) gauge and ghost fields and their Wick powers. Defining a gauge variation for a Wick monomial F according to

$$d_Q F \overset{def}{=} QF - (-1)^{n_F} FQ, \tag{155}$$

where n_F is the number of ghost fields in F, Q becomes a differential operator in the sense of homological algebra, and we have

$$d_Q^2 = 0 \Longleftrightarrow \{Q, [Q, F_b]\} = [Q, \{Q, F_f\}] = 0, \tag{156}$$

where F_b is a bosonic and F_f a fermionic operator and $d_Q(FG) = (d_Q F)G + (-1)^{n_F} F d_Q G$. For example, we get

$$d_Q : A_{\mu a} u_b \partial^\mu \tilde{u}_c :=: [Q, A_{\mu a}] u_b \partial^\mu \tilde{u}_c : + : A_{\mu a} [Q, u_b] \partial^\mu \tilde{u}_c : - : A_{\mu a} u_b \{Q, \partial^\mu \tilde{u}_c\} : \quad . \tag{157}$$

If $F = d_Q G$, then F is called a coboundary. The term (85)

$$\tilde{f}_{abc} : \partial_\nu A_a^\nu u_b \tilde{u}_c := \frac{i}{2} d_Q(\tilde{f}_{abc} : \tilde{u}_a u_b \tilde{u}_c :)\,. \tag{158}$$

is such a coboundary, which does not contribute to the physical QCD S-matrix.

Appendix D: Gauge invariance at n-th order

A short discussion is given here which explains how the simple condition of gauge invariance at n-th order in the coupling constant (46) emerges as a generalization from the first order condition (44).

A thorough mathematical treatment of the subject would involve a discussion of operator valued distributions, which is not given here for the sake of brevity. Nevertheless, all mathematical steps presented in the following can be put on a sound mathematical basis. It is useful to remind some basic facts: Free quantum fields like the free scalar field $\varphi(x)$ (9) are operator valued distributions on a Fock-Hilbert space, i.e. an operator $\varphi(g)$ is obtained after smearing out the field operator $\varphi(x)$ with a test function $g(x)$. This can be written *formally* as

$$\varphi(g) = \int d^4x\, \varphi(x) g(x), \tag{159}$$

where $g \in \mathcal{S}(\mathbf{R}^4)$ is in the Schwartz space of infinitely differentiable and rapidly decreasing test functions. Furthermore, the tensor product $\varphi(x)\psi(y)$ of free fields (and normally ordered products of free fields) is also an operator valued distribution, whereas the local product $\varphi(x)\varphi(x)$ is only defined after normal ordering. The 'value' of a distribution in a single point makes no sense in general. It is sometimes convenient to treat distributions like ordinary functions, as we will do it below, because in many cases the simplified insight gained from such a simplification allows to construct a full mathematical proof at a later stage.

As mentioned in the paper, the $T_n(x_1, ...x_n)$ are well-defined time-ordered products of the first order coupling $T_1(x)$, and they are expressed in terms of Wick monomials of free fields. The construction of the T_n requires some care: If the arguments $(x_1, ...x_n)$ are all time-ordered, i.e. if we have

$$x_1^0 > x_2^0 > ... > x_n^0, \tag{160}$$

then T_n is simply given by

$$T_n(x_1, ...x_n) = T_1(x_1)T_1(x_2)...T_1(x_n) \tag{161}$$

According to the definition (33), the $T_n(x_1, ...x_n)$ can be considered symmetric in $x_1, ...x_n$. Using this fact allows us in principle to obtain the operator-valued distribution T_n inductively everywhere except for the 'complete diagonal' $\Delta_n = \{x_1 = ... = x_n\}$ [47]. The construction is inductive since in the construction of T_n subdiagrams with lower order appear. If T_n were a C-number distribution, we could make it a well-defined distribution for all $x_1, ...x_n$ by extending the distribution from $\mathcal{R}^{4n}/\Delta_n$ to $\mathcal{R}^{4n}$. In the case of free field operators, the problem can be reduced to a C-number problem by the Wick expansion of the operator-valued distributions. The extension $T(x_1, ...x_n)$ is, of course, not unique: it is ambiguous up to distributions with local support Δ_n. This ambiguity can be further reduced by the help of symmetries (in particular gauge invariance) and power counting theory, and it is this local ambiguity which shows up as ultraviolet divergences in Feynman diagram calculations.

The concrete inductive construction of the T_n is discussed in detail in a famous paper of Epstein and Glaser [48] for scalar field theories, and explained in a more pedagogical way in [49]. For many practical calculations, it is advantageous to work in momentum space since the distributions, i.e. the Green's functions which occur in the calculations, behave much smoother in p-space than in x-space. The work of Epstein and Glaser is based on a treatment in real space, whereas in [49] practical calculations are performed in momentum space. An advantage of the real space formulation is also that it allows to formulate a consistent renormalization theory on a curved physical background [53]. Interesting topics like scale invariance, renormalization and the renormalization group can also be treated in real space [52, 50, 51].

As a specific example, we consider the second order contribution to the S-matrix which is given by

$$S_2 = \frac{1}{2!}\int d^4x\, d^4y\, T[T_1(x)T_1(y)] = \frac{1}{2}\int d^4x\, d^4y\, T_2(x, y). \tag{162}$$

We consider first the simple product $T_1(x)T_1(y)$ without time ordering. Then we have

$$[Q, T_1(x)T_1(y)] = [Q, T_1(x)]T_1(y) + T_1(x)[Q, T_1(y)]$$

$$= i\partial_\mu T^\mu_{1/1}(x)T_1(y) + iT_1(x)\partial_\mu T^\mu_{1/1}(y). \tag{163}$$

Considering the time-ordered product $T[T_1(x)T_1(y)]$, we have to distinguish three cases. In the first case, we have $x^0 > y^0$, and therefore $T[T_1(x)T_1(y)] = T_1(x)T_1(y)$. The second, a bit less trivial case is given for $x^0 < y^0$. Then we have $T[T_1(x)T_1(y)] = T_1(y)T_1(x)$, and

$$[Q, T(T_1(x)T_1(y))] = [Q, T_1(y)]T_1(x) + T_1(y)[Q, T_1(x)]$$
$$= i\partial_\mu T^\mu_{1/1}(y)T_1(x) + iT_1(y)\partial_\mu T^\mu_{1/1}(x)$$

$$= iT[\partial_\mu T^\mu_{1/1}(x)T_1(y)] + iT[T_1(x)\partial_\mu T^\mu_{1/1}(y)]. \tag{164}$$

Finally, we may have $x^0 = y^0$. If $(\vec{x} - \vec{y}) \neq \vec{0}$, i.e. if x and y are spacelike separated, then we may perform 'a little' Lorentz transformation such that $x^0 > y^0$, and the discussion above applies. If the special case that $x = y$, gauge invariance is not trivially fulfilled. This is not astonishing, since this is the case on the diagonal $\Delta_2 = \{x = y\}$ where perturbation theory may fail to calculate time-ordered products of the first order interaction. Requiring that gauge invariance holds also for $x = y$ leads to Ward-Takahashi identities in QED, and to the so-called Slavnov-Taylor identities for QCD.

The simplest example for such an identity is obtained from the vacuum polarization contribution to the S marix at second order, which can be written in the form

$$T_2^{vp}(x, y) =: A_\mu(x)t^{\mu\nu}(x - y)A_\nu(y) :, \tag{165}$$

where $t^{\mu\nu}(x - y) = t^{\nu\mu}(x - y)$ is a C-number distribution. Operator gauge invariance requires

$$[Q, T_2^{vp}(x, y)] = i\partial_\mu u(x)t^{\mu\nu}(x - y)A_\nu(y) + iA_\mu(x)t^{\mu\nu}(x - y)\partial_\nu u(y)$$

$$= i\partial^x_\mu[u(x)t^{\mu\nu}(x - y)A_\nu(y)] + i\partial^y_\nu[A_\mu(x)t^{\mu\nu}(x - y)u(y)], \tag{166}$$

and therefore

$$\partial^x_\mu t^{\mu\nu}(x - y) = \partial^y_\nu t^{\mu\nu}(x - y) = 0. \tag{167}$$

In momentum space, we get the C-number identity

$$k_\mu \hat{t}^{\mu\nu}(k) = k_\nu \hat{t}^{\mu\nu}(k) = 0, \tag{168}$$

hence the vacuum polarization tensor is transversal.

Appendix E: The adiabatic limit

The perturbative expression (33) is problematic, because the time-ordered products T_n are operator valued distributions after regularization, and they have to be smeared out by test functions. In order to be more precise in a mathematical sense, we introduce a test function $g_0(x) \in \mathcal{S}(\mathbf{R}^4)$ with $g_0(0) = 1$ and replace expression (33) by

$$S = \mathbf{1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int d^4x_1 \dots d^4x_n \, T_n(x_1, \dots x_n)g_0(x_1)...g_0(x_n). \tag{169}$$

Here, g_0 acts as an infrared regulator, which switches off the long range part of the interaction in theories where massless fields are involved. E.g., in QED the emission of soft photons is switched by g_0, and as long as we do not perform a so-called adiabatic limit $g_0 \to 1$, the matrix elements of the S-matrix remain finite. One possibility to perform the adiabatic limit is by scaling the switching function $g_0(x)$, i.e. one replaces $g_0(x)$ by $g(x) = g_0(\epsilon x)$ and performs the limit $\epsilon \to 0$, such that g approaches everywhere the value 1. If the S-matrix is modified by a gauge transformation, operators which are divergences are added to the n-th order term T_n. Such a contribution can be written as

$$\int d^4x_1...d^4x_n \, \partial_\mu^{x_l} O^{\cdot\cdot\mu\cdot\cdot\cdot}(x_1,...,x_l,...x_n) g(x_1)...g(x_l)...g(x_n)$$

$$= - \int d^4x_1...d^4x_n \, O^{\cdot\cdot\mu\cdot\cdot\cdot}(x_1,...,x_l,...x_n) g(x_1)...\partial_\mu^{x_l} g(x_l)...g(x_n). \qquad (170)$$

In the adiabatic limit, the gradient $\partial_\mu^{x_l} g(x_l)$ vanishes. Unfortunately, this property of the scaling limit does not guarantee that the whole term (170) vanishes (see also [54, 55]).

The infrared problem is not really understood in QCD, and all proofs of unitarity which exist in the literature have to be taken with a grain of salt, because they are avoiding the discussion of infrared problems somehow.

A thorough perturbative approach to the construction of the *local* algebras of observables which avoids the adiabatic limit is given for QED in [43] and (under the assumption that there are no anomalies) for non-Abelian gauge theories in [19].

Acknowledgements

I thank Dirk Trautmann and Florian Weissbach for carefully reading the manuscript. This paper is based on talks held at the Institute for Theoretical Physics at the University of Trento, Italy, and the Center for Theoretical Physics at the CNRS in Marseille, France.

The work was supported by the Swiss National Science Foundation.

References

[1] R. F. Streater, A. S. Wightman, *CPT Spin and Statistics and All That*, W. A. Benjamin, NY, 1964.

[2] J. Schwinger, Phys. Rev. **128** (1962) 2425.

[3] E. Abdalla, M. Abdalla, K. Rothe, *Non-Perturbative Methods in 2 Dimensional Quantum Field Theory*, 2nd edition, World Scientific, 2001.

[4] A. Aste, U. Walther, G. Scharf, Nuovo Cim. **A111** (1998) 323.

[5] C. Adam, Czech. J. Phys. **48** (1998) 9.

[6] C. Becchi, A. Rouet, R. Stora, Annals Phys. **98** (1976) 287.

[7] I. V. Tyutin, *Lebedev Preprint* FIAN 39, unpublished

[8] Y. Aharonov, D. Bohm, Phys. Rev. **115** (1959) 485.

[9] E. Fermi, Rev. Mod. Phys. **4** (1932) 125.

[10] A. Aste, G. Scharf, M. Dütsch, J. Phys. **A31** (1998) 1563.

[11] W. Pauli, P. Jordan, Z. Phys. **47** (1928) 151.

[12] S. Gupta, Proc. Roy. Soc. **63A** (1950) 681.

[13] K. Bleuler, Helv. Phys. Acta **23** (1950) 567.

[14] T. Kugo, I. Ojima, Suppl. Prog. Theor. Phys. **66** (1979) 1.

[15] M. Dütsch, T. Hurth, F. Krahe, G. Scharf, Nuovo Cim. **A106**, (1993) 1029, **A107**, (1994) 375, **A108** (1994) 679, **A108** (1995) 737.

[16] J. C. Ward, Phys. Rev. **78** 182 (1950).

[17] A. V. Razumov, G. N. Rybkin, Nucl. Phys. **B332** (1990) 209.

[18] A. Aste, G. Scharf, Int. J. Mod. Phys. **A14** (1999) 3421.

[19] M. Dütsch, F.M. Boas, Rev. Math. Phys. **14** (2002) 977.

[20] T. Hurth, K. Skenderis, Nucl. Phys. **B541** (1999) 566.

[21] P.A. Grassi, T. Hurth, M. Steinhauser, Nucl. Phys. **B610** (2001) 215.

[22] G. Scharf, *Quantum Gauge Theories: A True Ghost Story*, Wiley, New York, 2001.

[23] G. 't Hooft, Nucl. Phys. **B33** (1971) 173.

[24] T. Cheng, L. Li, *Gauge Theory of Elementary Particle Physics*, Oxford Science Publications, NY, Oxford, 1982.

[25] M. Dütsch, G. Scharf, Ann. Phys. (Leipzig) **8** (1999) 8.

[26] A. Aste, G. Scharf, M. Dütsch, Annalen Phys. (Leipzig) **8** (1999) 389.

[27] A. Aste, G. Scharf, M. Dütsch, J. Phys. **A30** (1997) 5785.

[28] M. Dütsch, B. Schroer, J. Phys. **A33** (2000) 4317.

[29] G. Scharf, Nuovo Cim. **A112** (1999) 619.

[30] I. Schorn, Class. Quant. Grav. **14** (1997) 671.

[31] G. Scharf, M. Wellmann, Gen. Rel. Grav. **33** (2001) 553.

[32] J. N. Goldberg, Phys. Rev. **111** (1958) 315.

[33] V. I. Ogievetsky, V. I. Polubarinov, Ann. Phys. **35** (1965) 167.

[34] W. Wyss, Helv. Phys. Acta **38** (1965) 469.

[35] R. M. Wald, Phys. Rev. **D33** (1986) 3613.

[36] D. G. Boulware, S. Deser, Ann. Phys. **89** (1975) 193.

[37] N. Grillo, hep-th/9911118.

[38] N. Grillo, Class. Quant. Grav. **18** (2001) 141.

[39] N. Grillo, Annals Phys. **287** (2001) 153.

[40] T. Kugo, I. Ojima, Nucl. Phys. **B144** (1978) 234.

[41] A. H. Chamseddine, Phys. Lett. **B557** (2003) 247.

[42] D. R. Grigore, G. Scharf, hep-th/0404157.

[43] M. Dütsch, K. Fredenhagen, Comm. Math. Phys. 243 (2003) 275.

[44] A. V. Razumov, G. N. Rybkin, Nucl. Phys. **B332** (1990) 209.

[45] J. Bognar, *Indefinite Inner Product Spaces*, Springer, Berlin, 1974.

[46] F. Krahe, Acta Physica Polonica **B27** (1996) 2453.

[47] R. Stora, *Differential algebras in lagrangean field theory* (ETH lectures, Zürich, 1994).

[48] H. Epstein, V. Glaser, Ann. Inst. Poincaré **A19** (1973) 211.

[49] G. Scharf, *Finite quantum electrodynamics, the causal approach* (second edition, Springer, Berlin, Heidelberg, New York, 1995).

[50] D. Prange, J. Phys. **A32** (1999) 2225.

[51] D. Grigore, Ann. Phys. (Leipzig) **10** (2001) 473.

[52] Y. Zhang, Eur. Phys. J. **C31** (2003) 103.

[53] R. Brunetti, K. Fredenhagen, Commun. Math. Phys. **208** (2000) 623.

[54] A. Aste, hep-th/0104014.

[55] G. Scharf, Nuovo Cim. **A109** (1996) 1605.